DOCUMENT RESUME

ED 179 230 IR 007 900

AUTHOR McNamara, J. E.; Johnston, T. S.
TITLE A Model for Cost Effectiveness Analysis of CBE Assessment Strategies.
INSTITUTION Syracuse Univ., N.Y.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
PUB DATE Aug 79
NOTE 89p.; Contains some light print

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Cost Effectiveness; Costs; Decision Making; Educational Administration; *Educational Assessment; Educational Planning; Efficiency; Matrices; *Models; *Performance Based Education; Program Effectiveness; Resource Allocations; *Systems Analysis

ABSTRACT

The Cost-Effectiveness Analysis Model for Competency Based Education (CBE) Assessment Programs is designed to assist state and institutional decision-makers in the cost analysis of alternative assessment strategies in CBE programs. The general model indicates stages of decision-making related to the use of various assessment strategies within the CBE context. Data are initially collected and recorded in two separate data matrices--one for cost and one for effectiveness. The outcomes of each of these analyses are grouped in a data display for decision-making. Included in the described model are procedures for cost analysis, primary cost analysis, secondary cost analysis, primary effectiveness analysis, and secondary effectiveness analysis; a cost data matrix; an effectiveness data matrix; and a summative data display for decision-making. (CMV)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED179230

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A MODEL FOR COST EFFECTIVENESS ANALYSIS OF CBE ASSESSMENT STRATEGIES

J. E. McNAMARA

and

T. S. JOHNSTON

AUGUST, 1979

IR007900

INTRODUCTION

Background

In July of 1978, Dr. Don Ely of the School of Education and Dr. Evelyn Daniel of the School of Information Studies initiated research designed to develop alternative strategies for evaluating competencies of library/media personnel in competency-based programs. This work was funded through a grant from the Office of Education. Their efforts have been directed toward collecting, developing, and testing strategies for assessing library/media personnel. The major outcome of this work was to be an assessment model capable of providing maximum flexibility in the evaluation of student performance within a CBE program. Parsimony was deemed to be the major criterion for the model. That is, the model was developed to maximize valid assessment within limited program resources. A portion of the project then focused on how the Daniel/Ely model could help identify program costs and effectiveness.

The critical economic question in attempting to design and improve programs is how do programs allocate their resources in the process of achieving their goals? This question is central to most program development. A second important question relates to program effect. That is, how effectively are programs using their available resources in achieving their objectives? Groteleuchen and Others (1976) address this question quite succinctly, asking, "What do you

get for what you give?" The bottom line, when these questions are answered is, is the program effective enough to warrant the resources allocated to it?

It was presumed that cost-effectiveness data would aid in choosing an appropriate mix of competency assessment strategies,* given a program's resources. It is obvious that costs and effectiveness can best be approached through cost-effectiveness analysis. This paper, then, provides a cost-effectiveness analysis model for CBE assessment strategies. Such a narrow focus, on assessment strategies only, means that the CE model presented here must preclude an examination of the broader competency-based setting. Thus, competency development, criteria and standards development, and instructional effectiveness will not be addressed. However, although the CE model to be presented below is referenced to CBE assessment, the authors believe that this model can be used to examine most typical evaluation procedures, whether CBE or not.

Audience

Cost effectiveness analysis is viewed in this paper as a form of evaluation which follows Stufflebeam's definition of

---

* The term "assessment strategies" has been adopted by ELY & Daniel and will be used throughout this paper. Synonomous terms in other contexts might be assessment/evaluation techniques, procedures, or methods.

evaluation for decision making: "Evaluation is the process of delineating, obtaining and providing useful information for judging decision alternatives" (1971, p. xxv). This paper, therefore, is primarily intended for decision makers concerned with evaluating various strategies for assessing competence. Typically, such decisions are made by either state governments or academic institutions. The decision makers, then, are members of state agencies (certification boards, departments of PSED, program accreditation teams, etc.) or academic administrators (academic deans, department chairpersons, etc.). The state agents often represent the public interest. For example, the Bureau Chief of the New York State Bureau of Libraries and Learning Resources is responsible for media specialist accreditation standards. State agents are likely to encourage academic programs which increase and/or claim to certify professional competence. A claim of competence, if justified, protects the public from malpractice in the professions due to lack of initial proper training. Thus, a bureau chief or state certification board can represent the interest of the public by deciding to support CBE and require programs to be competency based, or at least require adequate measures of competence at the end of a program.

Institutional concerns are generally the province of academic deans and department chairpersons. These decision makers represent the institutional interests, in particular, the development of quality programs and the optimal use of resources through administrative decisions, deans must choose

whether to promote and foster training programs which claim to teach competencies or to remain with more traditional degree structures. Chairpersons and instructors should work closely with deans in developing and implementing these programs. This is a critical link, because it is the individuals most immediately involved in instruction who can judge accurately whether CBE is a feasible academic option. The dean, on the other hand, can better estimate the financial resources available for such a program.

The cost effectiveness model outlined in this paper has been designed to respond to the needs of state and institutional decision makers, as well as to address the assessment problems associated with CBE programs. It is hoped, however, that the model will gain the attention of other educators in non-CBE programs, and that the model may be adapted to a more general class of cost effectiveness problems in student assessment.

Rationale for the Present Effort

CBE literature is replete with references regarding the difficulty of determining competency through performance assessment (Burke and Others, 1973; McDonald, 1974; Harris and Kelly, 1976; Reilly and Others, 1977; Madaus and Airasian, 1977). One solution often suggested for overcoming problems is the use of multiple assessment strategies. Hodgkinson and Others (1975) feel that the use of multiple measures permits

"triangulation" among different measures. That is, by comparing measurements of the same attribute mode with different instruments, it is possible to discover how well the attribute is defined and understood by how much the measurements overlap. Daniel and Ely (1978), in addition to acknowledging the value of the work of Hodgkinson and his collegues, argue that a wide range of alternative assessment strategies also provides an important dimension of flexibility to performance assessment programs. However, the use of multiple strategies creates a need for a method to choose an optimum mix from among many alternatives. Thus, the solution entails a further problem.

It is unlikely that there exists anywhere a single assessment strategy which is capable of measuring all the performances contained in a single competency-based program. The need for different assessment strategies seems inherent in the innovation. For example, the competencies identified and validated for the library media specialist (see appendix A) are so diverse that they must be assessed with different techniques. Still, selecting alternatives which provide credible evidence of competence is difficult and calls for a procedure by which one can select, not only the alternatives which measure competencies most accurately, but alternatives which are consistent with a program's content and available resources.

Cost-effectiveness (C-E) analysis is one solution for achieving this goal of selecting assessment alternatives.

This seems self-evident because the procedure covers the crucial elements involved in any realistic decision process; costs and effectiveness.

Cost: Hodgkinson and Others (1975) point out that competency-based education places new demands on student assessments. These additional demands entail additional costs, particularly, the continued cost of expanded operation of assessment from year to year. It has become necessary for CBE educators to consider these costs very carefully. Many CBE projects fail either because the financial resources and professional committment to operate CBE assessment are not available, is consistent with the mood of the times. The current rage for accountability necessitates a justification of program expenses. Before any such accounting can take place, however, it is necessary to determine where costs occur. The cost analysis model presented below is a means for getting at such information.

Effectiveness: Effectiveness, in this case, describes how good assessment outcomes are in judging future professional competence. "Good" here includes asking, "Are they believable?" and "Do they truly discriminate compentent from incompetent performance?" Since outcomes which are dubious reflections of performance tend to be disregarded, CBE programs based on such indicators are decidedly illfated.

The model and matrices which follow attempt to focus on these criteria, cost and effectiveness, and integrate them in a comprehensive manner.

Some Limitations of the Model

The costing procedures outlined in this paper deal soley with operational costs of assessment in CBE programs. In other words, the procedures presented below will provide a projection of typical yearly costs of CBE assessment per course, per pupil for different assessment strategies or mixes of strategies. However, there is a need to identify the other costs involved in CBE in order to provide a context for the discussion of operational costs, and to alert potential decision makers to the fact that many more resources are required to actually start a program.

For any new program, there are startup costs: those costs for planning, research, publicity, negotiation, and persuasion which go into formulating ideas and gaining acceptance for an innovation. Once a decision has been reached to institute a new program, there are development costs for, among other things, creating new or adapting old courses, instructional materials, facilities and tests. And finally, as mentioned above, there are operational costs associated with keeping the program running.

Thus, by considering only operational costs, the authors have made an implicit assumption that program decision makers have already documented the need to improve their curriculum, and have a sense of commitment to some new development effort. (In the case of CBE, a critical question to answer is whether

a validated set of competencies already exists, since competency generation and validation is a long and costly process). Once these prerequisites have been fulfilled, the question arises, "Should our effort be CBE?" The cost and effectiveness matrices in this paper are designed to collect and array data which will help clarify this decision.

CBE: A Definition

Competency based education originated as an improvement upon standard teacher education and training programs. It has spread to other professional education and training programs in the post secondary realm, notably, nursing, social work, library/media service and pharmacology. CBE as a curriculum innovation attempts to develop skills and attitudes which are essential for performing a set of specified tasks which, in turn, have been identified as necessary for success in a given profession. The move to CBE from other programs in PSED has been percipitated by the suspicion, and in some cases evidence, that specialized degree programs too often have failed to produce professionals who have the necessary skills to function at a minimum level of proficiency when they are placed in the field. In other words, traditional advanced degree training is no guarantee that a person can perform adequately in the job for which he/she has prepared. The advent of CBE is an attempt to remedy this discrepancy.

Differences Between Graduate CBE and Lower Levels

CBE is currently experiencing a great popularity among state secondary education programs, and consequently, general definitions and descriptions of CBE tend to focus on this level. The heart of any CBE program is that any person with an academic credential is certified as having demonstrated competence in certain skills. While this central tenent is found in graduate CBE, many other assumptions, which accompany CBE at lower levels, are not. Some of these differences are critical, especially from the perspective of cost effectiveness analysis. Several important distinctions will be elaborated below.

The major distinction between graduate and non-graduate CBE is that the model of mastery learning is not necessarily assumed at the graduate level. Mastery learning asserts that the majority of students can learn the contents of a given curriculum if they are provided enough time to master the material at their own pace. In addition, in such programs, students are often tested when they feel that they have mastered a certain portion of material, rather than when the instructor feels that they should have completed that portion. According to this format, there should be virtually no failure in masterylearning, because students may retake (in some cases must retake) tests which they have not passed. They may not proceed beyond a given level of material until they demonstrate proficiency.

CBE mastery learning, by extension, is mastery learning where the material studied and the level of proficiency required are referenced to some externally validated set of performances and standards of achievement for some real-world skill. At the graduate level, these skills are most often professional performances related to specific professional roles.* Most graduate training programs and professional schools which have adopted CBE have not accepted the mastery model completely, for several reasons. First, there has already been a thorough screening and self-selection process which is designed to guarantee that those in graduate programs are capable students and can accomplish the work presented in the curriculum. On the other hand, because the levels of responsibility and expectation are much higher for graduate students, and the competition for places in graduate school is extreme, there is some sentiment that individuals should not be allowed an indefinite period to master material. Finally, the professionals being produced by graduate schools are placed in positions of great responsibility and influence. (Many medical people literally deal with life and death, while others, social workers and teachers, for example, can have lasting influence on clients lives and careers.) Therefore, it seems neither reasonable from a school's perspective, nor responsible

---

* There are broader categories of practical behaviors such as "survival skills" and "coping skills" which cover basic knowledge and behavior necessary to lead a minimally fruitful life. We will restrict our discussion to professional competencies necessary for job performance.

from the view of society to allow students an extended period to reach mastery, when repeated failure may indicate a lack of ability or disposition for professional practice. When a school's reputation and that of its graduates may be compromised in the public eye by what is seen as an over-lenient attempt to qualify inherently unqualified individuals, then competency education gives way to accountability. Thus, CBE in graduate school is less an attempt to deal fairly with each student in allowing that student to mature at his/her own pace, than an attempt to certify that each graduate is capable of delivering professional service in the field upon successful completion of a course of study.

A second aspect of CBE in secondary school, which does not appear often in graduate education, is the use of multiple approaches to a single learning module. A fundamental educational principle is that different individuals learn in different manners. Thus, CBE often presents several ways to learn the same material. Graduate students, however, have shown by their very presence in school, that they can learn, if not optimally at least satisfactorily in a traditional school setting. For this reason, as well as the added costs of development, and the more diverse demands on graduate faculty, there is generally only a single method of instruction available.

A third aspect of CBE which is neglected at higher educational levels is remediation. The cost of graduate

education (approximately $_____ greater than undergraduate education) is so great, especially in an era of scarce educational funds, that the remedial element of competency education simply is not feasible. The time lines for traditional academic programs are simply not flexible enough for complete CBE. Neither are the opportunity costs to students, schools, and society justified in allowing one ill-fitted student to use resources which might train one or more better candidates.

Finally, CBE at the graduate level is generally more idealistic than at the secondary or primary levels. The CBE movement at lower levels of education is designed to guarantee that individuals have the minimal skills to survive in modern society. While these skills are critical for those who must attain them, they are necessarily a lowest common denominator for possible curricula. They are also directly referenced to specific daily activities. The competencies in graduate school, however, are demonstrated and required more to protect the clients of the graduate professionals than to assure students of personal well-being or fulfillment. The competencies of graduate training are also designed to be both skills necessary at present in the field, documented through field practice, as well as competencies which ought to be used, but because of a lack of training or opportunity, are not being demonstrated in current practice. Additionally, competencies which are projected as necessary in the future are also often included in graduate courses, while such potential skills are not so crucial

in less advanced education.

These differences between graduate/post secondary CBE and lower levels are critical in planning and executing a curriculum. While losing many of the desirable features of competency education, graduate programs can better concentrate on the critical feature of CBE: assessment. It is to that aspect of CBE that the paper now turns.

Assessment of the Centerpiece of CBE

There are three basic assumptions that underlie CBE (Cyrs, 1976). First, for most jobs, it is possible to identify a finite group of observable behaviors, (competencies) based on specific skills and attitudes which are necessary and sufficient for competent professional behavior. Second, appropriate performance measures for each competency or group of competencies can be designed and validated. Third, in order for a program to claim with the most possible confidence that an individual has actually mastered a minimum level of necessary proficiency to function as a capable professional, performance assessment should take place in the actual professional settings.

It is evident from these three assumptions that evaluation is central to CBE, particularly if a validated set of competencies has not been previously identified. In fact, Harris & Kelly (1976) argue that CBE "cannot succeed unless the profession (of educators and measurement specialists) emphasizes

the development of performance assessment strategies and *techniques much more,* (and that) ...the credibility of any competency-based curriculum rests directly on the adequacy and appropriateness of its assessment procedures" (pp. 241, 247). Without thorough assessment, there is little reason to assert that students taught in a CBE curriculum are any more capable of performing well on the job than those taught in regular curricula. Thorough assessment means both continuous and redundant measurement of students in the field, on the job, or in job-like settings. Assessment must be continuous in order to provide feedback to students about their performances, expecially when they seem sub-standard, and redundant in order to provide multiple measures of the same behavior (hopefully, with different instruments) so that students are evaluated with precision and consistency.

The Central Tradeoff: The Cost vs The Validity of Assessment

Because thorough assessment is so fundamental to CBE, it would be ideal to try to use the most sophisticated, individualized, and accurate assessment strategies which could be devised. This would guarantee to students that their training had indeed been effective and worthwhile, and to the public that graduates are competent. However, it is generally assumed in CBE that reliability and validity of performance assessment vary directly with costs. That is, if two assessment strategies are equally well constructed, the strategy that best measures com-

petent performance will be more expensive. Thus, the strategies listed in this paper as potential measures of competence become more expensive as they measure competence more thoroughly. However, though validity is a good predictor of cost, cost is not necessarily a predictor of validity. In other words, if a measure is more valid for measuring performance, it cost more, but simply because a measure is expensive does not mean that it measures well.

The real difference in assessment strategies, then, is how well they determine the achievement of program competencies. For example, even the most well-designed paper and pencil test of class-based knowledge, the type of test most pervasive in education, generally assesses only a student's cognitive abilities, and perhaps some attitudes. Paper and pencil measures are relatively inexpensive. There is a critical assumption, however, that these cognitive measures correlate highly with competent performance in an actual job-setting. Where this assumption is greatly in error, as it is in most professions, (Pottinger, 1975) the paper and pencil test is a poor measure for CBE. At the other end of the spectrum of assessment strategies, continuous observation and feedback of every competency, performed in actual job situations, over an extended period of time virtually guarantees that an individual can perform certain skills and exhibits certain attitudes at requiste levels of proficiency. Furthermore, there is more reason to assume that these behaviors will transfer to other job settings. However, this kind of

competency performance assessment is probably beyond the scope of any actual job performance assessment done by employers or any training assessment done in schools--except perhaps the training of pilots, physicians, nuclear technicians and other individuals whose jobs involve narrow margins of error between life and death. One major reason that few professions use such rigorous performance assessment techniques is that the costs of detailed assessment become prohibitive very quickly. Another reason is that the technology of performance assessment is much less well developed and articulated than that of paper and pencil tests and other non-performance measures.

Thus, the tradeoffs in CBE are pretty clear: more valid claims of competent job performance by graduates cost more money. Therefore, the critical questions for any group contemplating CBE as an alternative curriculum are these: What standards of validity are acceptable in our assessments? Can we afford the assessment strategies necessary to guarantee these standards (if they are not already there)? Is the cost of assessment worth the increased certainty of performance? Is the cost differential between various strategies worth the incremental validity produced by different strategies? What is the purpose of our program: to prepare people who can learn in the field, but who have not necessarily mastered all aspects of the profession? Or to produce graduates who are at least minimally competent professionals upon completion of the program?

Assessment is Assessment is Assessment

There is a basic clarification which should be made at this juncture between the costs of assessment in CBE, and the costs of good assessment in general. The difference between assessment strategies which might be used in CBE and in non-CBE curricula does not become pronounced until the strategies become demonstration based. In the three demonstration modes which will be presented below, a significant increase in resources, especially in teacher time required to learn observational techniques, to observe students, and to aggregate and evaluate data, is necessary for valid and reliable assessment. If these strategies are not available, then performance assessment in demonstration strategies is not worthwhile, and consequentially, the more valid ranges of CBE assessment must be foregone. Stated differently, no assessment strategy will provide valid and reliable data if it is done inadequately. And, the lack of resources is almost a sure guarantee of inadequate assessment.

If a program has already invested in well-developed and validated paper and pencil tests which teachers feel, with some degree of confidence, are good indicators of performance, then it makes little sense to abandon them for more sophisticated measures when the lack of resources means that the more sophisticated strategies cannot be implemented correctly. Well-conceived, and developed paper and pencil tests, then are better indicators of performance than poorly conducted observations.

Thus, it should not be assumed that techniques which are directly performance-based are inevitably better than indirect indicators. Rather, the criterion for increasing validity and reliability is that the appropriate initial level of investment must be made for any mode of evaluation to be effective. Only after the investment is made to guarantee that all modes of evaluation are initially well-developed can a comparison of one mode with another yield an index of relative costs and effectiveness.

It is critical to realize that the development of good assessment techniques and instruments, no matter in what mode, is a costly process. And, even with the most common kinds of tests: traditional paper and pencil objective and essay exams, development of valid and reliable tests is often a complex, multi-stage process. Therefore, development of evaluation tools and procedures, whether in CBE or traditional curricula, is a substantial undertaking. Costing CBE, then, should be considered in the context of the notion of a general improvement of evaluation procedures. If a program needs improvement, it is likely that evaluation procedures will need to be revamped or totally redeveloped. The issue, then, is not really in the costs of redevelopment, since this would occur in any new program whether it were CBE or not, but in the costs of operating the evaluation procedures. The level of validity and reliability necessary to make claims of competent performance are higher than to make claims about cognitive knowledge gains. Therefore,

CBE generally requires most costly operations to reach these higher levels. To do good evaluation, it is necessary to revise and update procedures and measures, even for the most basic kinds of paper and pencil tests. Thus, the question is not whether to use CBE or not, but how to get the best evaluation procedures and techniques per each dollar spent, that are congruent with the goals of the program.

If a program cannot presently afford to maintain even a modest evaluation program, then CBE certainly is not an option. If a non-CBE program currently has a set of good evaluation procedures, and a strong program then a more costly set of good CBE evaluation techniques is only worthwhile if the goals of the program shift to encompass CBE criteria, and if costs are not a critical factor.

There are two other cases which should be considered. First, program personnel may presently have a good evaluation system which does not measure performance, but they perceive a need to improve the curriculum. If the program administrators are considering CBE, they must decide whether they can afford effective performance measures and procedures. (There will be significantly marginally greater development costs if a new evaluation system as well as a new curriculum must be developed.) If not, then the present system will probably be inadequate to make claims of competence for graduates. The second instance to be considered is a program which begins with a poor eval-

uation system (initially). If both the curriculum and assessment procedures are to be changed, then the differential costs between developing valid new paper and pencil measures, for instance, and developing simulations for performance assessment will not be as great as when there are already good measures in place. Thus, when CBE is an option, the cost of operation of assessment is the major consideration. In summary, then, CBE should be considered when program personnel are convinced that: 1) the desire and resources are available for change, and 2) the desire and resources are available for maintenance of a strong assessment program. Once these prerequisites have been satisfied, the actual feasibility of operating a CBE assessment system can be estimated with the cost and effectiveness procedures outlined in the remainder of the paper.

COST-EFFECTIVENESS MODEL

This section discusses the Cost-Effectiveness Model for analyzing assessment programs in competency-based educational programs. The overall Model is discussed first, followed by a detailed discussion of 1) the cost analysis activities within the model, and 2) the effectiveness analysis activities within the model. The section concludes with a discussion of the matrices employed in the analyses. The purpose of this section is to give the reader an overview of the cost-effectiveness analysis being conducted on the CBE assessment program.

The General Model

The general cost-effectiveness model for CBE assessment programs outlines the activities around which the cost-effectiveness analysis is conducted. The model is designed to be comprehensive without being excessively elaborate and confusing.

Figure 1 displays the general model.

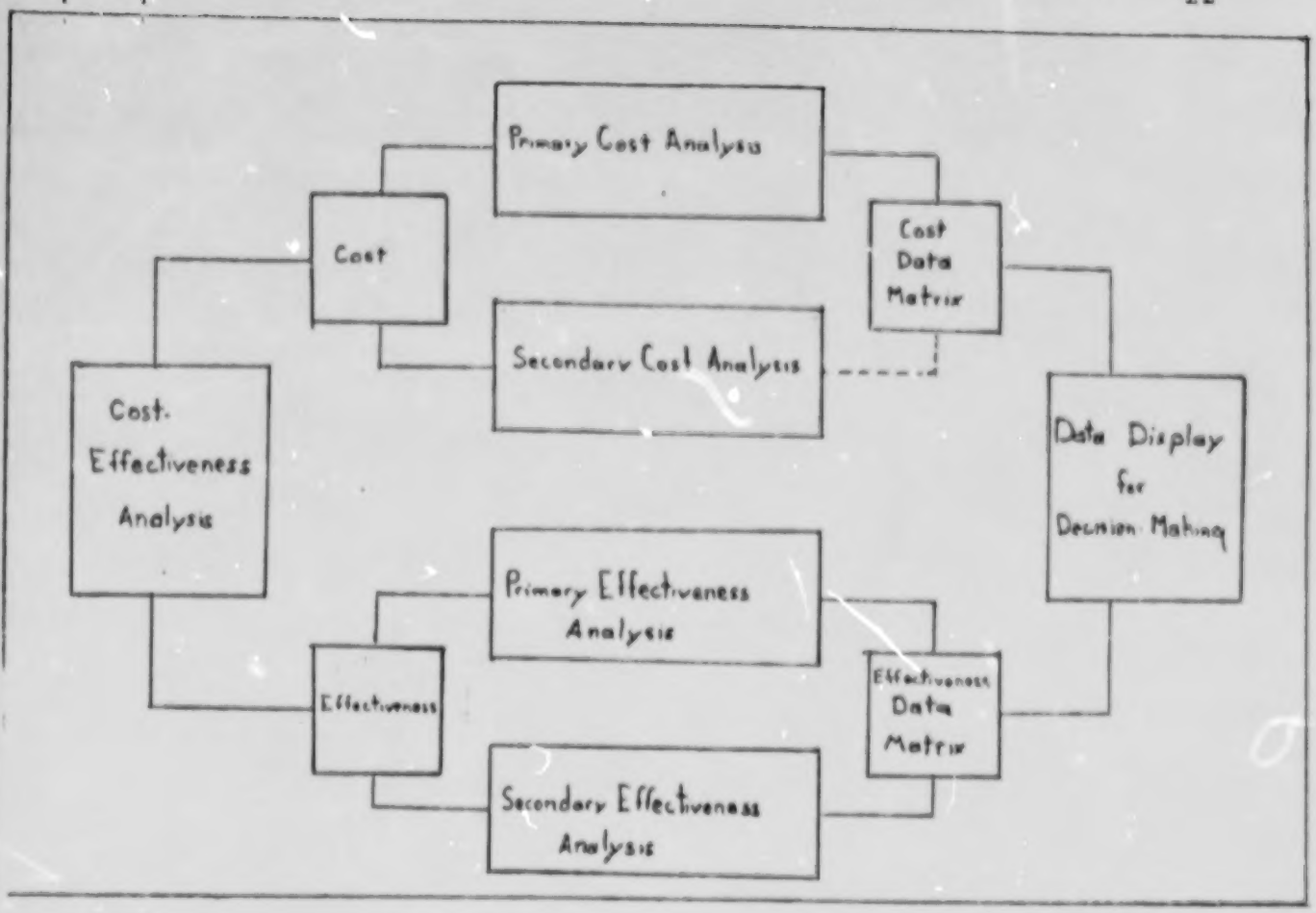


Figure 1. General Cost-Effectiveness Analysis Model for CBE Assessment Programs

The model is drawn to illustrate the way in which the cost-effectiveness analysis of assessment programs is to be approached: that is once as a cost analysis, and again as an effectiveness analysis. The reader should keep in mind that the analysis is conducted for the purpose of exposing differences between alternative assessment strategies such that these differences assist in the policy development relating to the assessment program. The bulk of the analysis effort is concentrated in the center of the model and is prepresented by the four blocks:

1. Primary Cost Analysis
2. Secondary Cost Analysis
3. Primary Effectiveness Analysis
4. Secondary Effectiveness Analysis.

Initially the results of cost and effectiveness analysis are placed within two separate data matrices:

1. The Cost Data Matrix
2. The Effectiveness Data Matrix.

Ultimately, however, the data from these matrices are brought together in the Data Display for Decision-Making.

In the paragraphs immediately following the various analyses are discussed in the most general manner. This discussion is intended to prepare the reader for the more detailed discussion of 1) cost analysis in the section entitled "Cost Analysis", and 2) effectiveness analysis in the section entitled "Effectiveness Analysis". It is the authors' goal to provide a basic, although thorough, presentation of cost-effectiveness for CBE assessment programs.

Cost Analysis

In general cost analysis attempts to discover the potential and/or actual costs of a program. As the term is used here, cost means the expenditures of human and monetary resources for the development and operation of programs. As it is currently conducted cost analysis concentrates largely on the monetary resource expenditures. However, there is

increasing attention given to the expenditure of human resources as well (e.g. Schalock, In Press). This present distinction of human and monetary costs provides the basis for a two-part analysis of costs; namely: Primary Cost Analysis and Secondary Cost Analysis.

Primary Cost Analysis

In general the primary cost analysis follows the design laid out by Doughty and Others (1976). This cost analysis examines the major cost categories of resources (material and monetary). "These categories serve to specify the segments that require analysis" (p. 27). Program activities are analyzed in terms of the expenses they generate along the cost categories of resources, or Activities X Resources.

Doughty and Others (1976) also examine development and operation costs of programs. That is, they attend to the expenditure of resources as these relate to both the development of a program and to its operation. Recall, however, that the present model concentrates only on operational costs. That is, the general C-E model under discussion here disregards developmental costs.

In this model the two dimensional analysis of Doughty and others (1976) is conducted along another dimension,--that dimension includes the eight assessment strategies. This results in a three dimensional cost analysis; one which examines

the costs generated by the interactions of Resources with Activities with Assessment Strategies; or

Resources X Activities X Strategies

This cost analysis model is discussed in detail below.

Secondary Cost Analysis

By and large current cost analysis disregards non-material costs because it is difficult to assign numerical (or dollar) values to these costs. Secondary cost analysis, however, consciously considers these non-material costs. Typically these are social costs, psychological costs, lost opportunity costs and the like. Because dollar values are not easily assigned to such costs, analysis tends to overlook them, and in turn discounts important decision-related information.

Primary Effectiveness Analysis

The primary effectiveness analysis has been designed to deal only with the assessment outcomes. This analysis asks the question: To what degree do the assessment strategies accomplish what they set out to accomplish? That is, how effective are they? Put in this way, the effectiveness analysis of CBE programs takes on a unique characteristic particular to the field of Tests & Measurement. For in asking whether or not an assessment strategy accomplishes

what it set out to accomplish, is the same as asking: Is the assessment strategy valid? In this sense effectiveness analysis of assessment systems is the same as validity analysis. The effectiveness analysis, therefore, will by and large be concerned with determining the validity of the various assessment strategies. The format for this analysis is laid out in detail below.

Secondary Effectiveness Analysis

Validity analysis of assessment strategies has become such a sophisticated technology that it resists certain secondary forms of analysis. That is, validity analysis relies upon certain test and measurement operations. These operations are inappropriate when considering questions relating to 1) the practicality of an assessment approach, 2) the complexity of directions within an assessment instrument, 3) the length of a given assessment instrument and so forth. An appropriate blend of these types of questions and the validity question has yet to be developed. Secondary effectiveness analysis provides important decision-related information and analysts should exercise caution in overlooking secondary effectiveness analysis.

The Value of a Dichotomous Approach to Analysis

The dichotomous approach to analysis is undertaken for

the sake of completeness. It has the advantage of pointing up to the analysts what they typically overlook in approaching C-E analysis. This illustration of primary and secondary parts of analysis gives analysts an opportunity to see what is critical and what is not so critical and subsequently permits a choice for analysis.

Although complete cost-effectiveness analysis requires both primary and secondary forms of analysis, circumstances may dictate against both forms of analysis. Evaluation and decision-makers, therefore, will have to make choices about which elements of secondary analysis will stay in the analysis and which will go out of it, and in doing so they will need to be aware of the reasons why certain secondary analysis has not occurred.

The Matrices

Cost analysis and effectiveness analysis are greatly facilitated by the use of matrices. Doughty and Stakenas (1973) make a strong recommendation for their use in reporting and relating information for decision-makers. In addition to their value in reporting and relating information, however, matrices are useful as a format in which to arrange data en route through the analysis. Often matrices become part and parcel of the analyses process and may even serve to guide the analysis itself. Matrices are used in the present model for

all of these reasons. This section briefly discusses the matrices which are used in the model. There are three: 1) the Cost Data Matrix, 2) the Effectiveness Data Matrix, and 3) the cost-effectiveness data matrix referred to here as the Data Display for Decision-Making. Each matrix is discussed with respect to the way in which it is used within the C-E analysis.

The Cost Data Matrix

The cost data matrix is used for several reasons: 1) to guide cost analysis; 2) to display and relate information for decision makers, and 3) as a residual for the placement of available data. For the present the cost Data Matrix is confined largely to primary cost analysis. The growth and development of the cost-analysis of competency-based assessment programs may expand the use of this matrix to Secondary Cost analysis as well.

The Cost Data Matrix serves to display the costs for each strategy analyzed and to prepare the cost data for comparison with the effectiveness data. An illustration of this matrix is provided on p. 49.

The Effectiveness Data Matrix

The effectiveness data matrix is used 1) as a residual for effectiveness data and as a reporting device. Principally the Effectiveness Data Matrix serves to display effectiveness

data for analysis and interpretation. Unlike the cost matrix however, the effectiveness matrix is not confined to primary analysis. However, the degree to which it will include secondary forms of effectiveness analysis data is highly situational. Because an effectiveness analysis is essentially a validity analysis, the data form such an analysis may not require a matrix. The effectiveness data matrix is recommended, nevertheless, because 1) its use indicates to the evaluator the breadth of the analysis and 2) forces the evaluator to ask secondary analysis questions.

Data Display for Decision-Making

The final major element in the general model is the Data Display for Decision-Making. For the display of analysis data Doughty and Stakenas (1973) suggested a tabular array approach. Their claim is that such an approach

> enables comparisons to be made of different kinds of quantitative data as well as qualitative criteria. Single dollar cost measures of an alternative are as uncommunicative and uninformative as any contrived single measure of effectiveness. (p. 182)

They emphasize the need to let the array "reflect a set of cost and effectiveness measures suitable for use in making decisions" (Ibid). They further suggest a two dimensional field for comparison, which in this case is Strategy by comparative Cost Effectiveness criteria or

Assessment Strategy X C-E Data

Figure 2 illustrates this arrangement.

Figure 2. Data Display for Decision-Making

Strategies / Criteria	Field Experience	Actual Demonstration	Simulated Demonstration	Portfolio	Papers-Projects	Oral Presentations	Oral Exams	Written Exams
Effectiveness:								
Primary								
Secondary								
Cost:								
Primary								
Secondary								

Doughty and Stakenas state the advantages of this arrangement.

> The development of the matrix provide(s) for systematically organizing, analysing and synthesizing the data. The matrix can display data that are accessible to examination so that it can be easily revised and modified as need and additional information becomes available. Data comparisons are possible within an alternative just as easily as they are between alternatives. Both types of comparisons are necessary if proper analysis is to be conducted.
>
> (Ibid)

Despite its advantages, the matrix does not make decisions for the decision-maker. At best it merely informs the decision-making process. Decisions will need defensible rationales and the data thus accumulated lends substance to these decisions.

Summary

This section has presented the General Cost-Model. Each major element within that model was briefly discussed. Emphasis was made with respect to the model's orientation to the decision-making process. The point was made that neither the model nor the final display make (assessment) program decisions. Decision-makers are needed for that. The next section provides a more detailed discussion of two analysis effects: 1) the cost analysis, and 2) the effectiveness analysis. The discussion will outline how these analyses operate specifically within the model.

The Cost Analysis Model

This section outlines the model used in the cost analysis of alternative assessment strategies in competency-based programs. The model contains two forms of analysis: Primary Cost Analysis and Secondary Cost Analysis. The necessity for two analysis procedures has been discussed in the section covering the General C-E Model. Although each form of analysis is discussed below, the major emphasis is on primary procedures of cost analysis.

Function/Resource Cost Analysis: Rationale and Description

Rationale: The purpose of cost-effectiveness analysis is the comparison of alternatives (programs, assessment strategies, etc.) in order to choose among them. The dimensions of cost and effectiveness are usually the critical variables in any such decision. Central to the very concept of comparison is the necessity of having some universal standard for these dimensions, and a good sense of what goes into and comes out of a given program or activity. In cost analysis this universal standard is dollars (rubles, pesos, dracmas, yen, etc.). There are difficulties with this standard which will be mentioned below, but at present, it seems the best reflection of cost available. Cost, as defined by economists, is "the best alternative forgone". (Fisher, 1975, p. 25.) However for this paper, cost will be used in a more colloquial

sense stated by Carpenter (1978) as, "a common and universal measure of the nature and quantities of resources used in a specific manner to achieve stated objectives" (p. 8). The stated objective here is the most valid measurement of competencies given the resource constraints of a (potential) CBE program.

Cost analysis in education must be function/resource based if it is to be helpful (Doughty & Beilby, 1974; Carpenter, 1978; Stace, 1979). "Function/Resource" is the authors composite term for costing procedures which are known both as Functional cost analysis and Resource cost analysis. Since the procedure necessitates the identification of functions and then their costing by resources used per function, the joint term seemed appropriate. Any time the reader sees either Functional or Resource in reference to cost analysis, this may be read as synonomous with or equivalent to function/resource. Methods of costing which are not based on the resources necessary for each individual, for specific program or activity can suffer from one of two possible errors in estimation. First, costs may be included which should not be charged to any single program, because, for instance, they are shared among several programs. These are referred to as joint costs, and will be discussed more fully later. Including joint costs entirely in a single program or activity thus inflates the apparent price. The second error mirrors the first; the analysts omit costs which a program incurs, but which are obscured by

more traditional accounting and budgeting procedures. This error depresses the actual cost of a program.

Carpenter summarizes the function/resource approach to educational costing very succinctly:

> (Function/) Resource cost analysis is designed to provide cost estimates which permit equitable comparisons among projects (strategies), where cost or cost/effectiveness is the basis of comparison... The process begins with determining the resources used in an institutional project. The cost of these resources - staff, facilities, equipment and materials - is then estimated using standard prices for comparable resources (p. 8).

Identifying resources which are used solely in a program or, in this case, an assessment strategy or system, however, is generally not simply a matter of going to a school budget official and asking for the various costs of resources in a program. This is because institutional budgeting and accounting systems are generally not designed for nor useful in planning or evaluation. Traditional budgeting and accounting are based on jurisdiction rather than function/resource. In function/resource budgeting, costs are distributed in relation to the programs in which they occur. Thus, costs are identified according to the activities of teachers, administrators, and students, and the resources used in those activities. This method of costing is in marked contrast to jurisdictional, line-item accounting which is used for most budgeting and auditing purposes in educational organizations. Jurisdictional

accounting, as its name implies, is based on the organizational unit to which funds are allocated, and not on the actual functions for which those funds are used. This procedure has two major drawbacks in attempting to determine how much a given program, or set of assessment procedures, will cost. First, jurisdictional costs are too general, and second, they are too undifferentiated. An example will help make the advantages of function/resource costing clear.

If, for instance, a dean wished to compare the costs of two programs for teaching Freshman English, s/he would need to know what activities (functions) were required for each program, and how much time and what materials (resources) were required for each activity. Traditionally a faculty member's salary is allocated to a department as part of an undifferentiated line-item called "professional staff". In addition, within the department, programs are usually not specifically broken out in terms of their use of resources. This is particularly true where faculty time may be put into inter-departmental programs or similar joint endeavors which belong to no department(s) separately. If a faculty member who is teaching in the Freshman English program is also teaching other courses, advising, performing research, and serving on committees and engaging in other school administrative matters, it would be misleading to assign the entire faculty time (salary) to the Freshman program. Instead, some estimate of the percentage of time spent on the Freshman program must be obtained, and trans-

lated into a percentage of faculty salary, including benefits. This figure, then, would be allotted to the particular program. The jurisdictional line-item budget usually cannot provide this kind of detail and differentiation. Thus, comparison of alternatives requires a clean distinction between line-item costing and function/resource costing, if the comparison is to be accurate and meaningful.

Description of the Cost Analysis Process: The general method for function/resource costing has been sketched above. There are several additional comments, and particularly, some "Precautions and Admonitions" which need further elaboration. To reiterate, in order to identify the resources which go into a specific program, it is necessary to identify the various functions which are included in that program, and thus, to specify what costs are constituted by, for example, "professional staff" in a given program. Costing generally occurs, then, by estimating the percent of time for faculty and staff, the percent of use for facilities, and equipment, (where not totally allocated to a single program), and the cost of consumables. (Typically for education, the major costs are for faculty time.) These functional estimates are then translated into dollar costs, where possible, from budget figures.

Two fundamental requirements for formulating functions and resources are that these categories must be delineated and defined so that they are explicit in identifying behaviors and

and activities, and they have to be mutually exclusive and exhaustive. These stipulations are particularly relevant for major resources such as personnel. If the functions and resources are not explicit, then an observer or a participant can not tell what is included when estimating different functions, or in what category a certain resource belongs. If they are not mutually exclusive and collectively exhaustive, function and resource categories may duplicate costs which appear in two functions simultaneously, or omit costs for which functions/resources have not been specified.

Once the functions and resources have been identified, they form the axis of the cost analysis matrix, (fig. 3, p.51). A third axis consists of whatever dimension is to be judged; in the current effort, this axis is made up of the various strategies.

Some Precautions and Admonitions: A cursory inspection of the cost analysis matrix makes the costing procedure appear deceptively easy. However, there are many major and minor difficulties which should be acknowledged in conducting cost analysis, and by extension cost-effectiveness analysis. Some of the ones of most concern for costing CBE assessment strategies will be mentioned below. However, for a fuller treatment, the reader is encouraged to consult the references cited in this section. Five cost-related concerns which are fundamental to the analysis described in this paper, and which

will be sketched below, are:

1. distinguishing different types of costs.
2. Fixing costs or effectiveness
3. differentiating strategy comparable costs from program specific costs
4. dealing with non-quantifiable costs
5. increasing the validity and reliability of time and usage estimates.

1. Distinguishing different types of costs: Costs can be categorized according to a variety of criteria. It is essential to be able to distinguish among these different cost categories, because confusing them with one another will distort the entire analysis. These distortions generally appear as over-estimates or under-estimates of costs. Such errors can radically affect the comparison among various alternatives where cost is a critical factor. Some of the basic costs which a cost analyst must be able to differentiate are: fixed and variable costs, recurring and non-recurring costs, joint costs, and sunk costs.

Fixed costs are those costs which remain stable despite fluctuations in a program's operations. Most major capital expenses are regarded as fixed costs. Costs for facilities, then, are usually fixed, since building mortgages and upkeep continue at a fixed level whether there are many or few students. (Obviously, this is true only within certain ranges; if there are no students, buildings may be leased or sold). Variable costs are costs which are directly related to the level of activity of the system. Thus, expenditures for materials are variable costs, since the more students there are, the more materials are required.

Recurring and non-recurring costs are self-evident, but extremely important. Recurring costs are those which occur during each and every cycle of a program's operation. Non-

recurring costs are one-time costs during a program's life, for instance, costs for research and development. If these costs are confused, especially when a non-recurring cost is listed as a recurring one, the entire program operation cost is inflated. In the cost matrix presented for use here, there are only recurring costs, since the analysis deals solely with operational costs of assessment systems.

Joint costs have already been mentioned. These are costs which must be allocated to more than one program or jurisdiction. They are costs which are shared, for example, different science programs' sharing the same equipment or teachers teaching in more than a single program. Determining percentages of use and time for such joint costs is crucial but difficult, and will be discussed more extensively under Reliability of Estimation, below.

Such costs are those expenditures which have already been made at some time in the past. A common error in cost analysis and in decision-making is to include sunk costs in future projections of the costs of a program. Once resources have been irretrievably allocated, they should be disregarded in calculating program costs. This reflects the approach of using the "present worth" concept of economic analysis, which allows all costs and decisions to be based on the value or resources at the moment of analysis. The general principle behind ignoring sunk costs is to avoid committing resources which are still flexible in the present, on the basis of the

past worth of resources which have been allocated and therefore, in effect, no longer exist. This procedure allows each decision to be based solely on the current value of flexible, available resources.

2. Fixing costs and effectiveness: Because the major purpose of cost-effectiveness analysis is the comparison of programs, strategies , or other like entities, it is usually necessary to fix at least one of the two standards of comparison: either cost or effectiveness. Fixing costs or effectiveness means setting some minimally acceptable level of effectiveness, or some maximally acceptable level of cost, and allowing the unfixed factor to vary. Thus, in comparing systems of assessment strategies, if there is a fixed resource limit beyond which costs cannot range, then the effectiveness of systems whose costs are under this ceiling is the criterion on which a selection will be made. Conversely, if a certain level of achievement is essential, for instance, reading at the national norm for grade level, then cost becomes the principal factor in choosing among programs which reach this achievement level or above.

There is another method for using cost and effectiveness criteria in which both cost and effectiveness are allowed to vary. In this method, alternatives are compared in pairs, the winner of each successive comparison being pitted against the next alternative until a single alternative remains. This

method has advantages over fixing either costs or effectiveness because it allows optimization along both variables rather than just one. However, there are technical difficulties in any multiple variable decision analysis which complicate the process. Among these difficulties is the eventuality that alternatives are superior on different criteria. In considering assessment strategies, this is nearly always the case, because as strategies become more effective in measuring competence, they become more expensive. Thus, decision-makers must devise some weightings of criteria to choose between such cases.

Another difficulty with not fixing either costs or effectiveness is what is known in political science as "the paradox of voting" (Riker & Ordeshook, 1976,p 34-5). Simply, the paradox of voting states that if there are at least three alternatives, a, b, and c, which are judged on several criteria, it is possible to prefer a to b, b to c, and c to a. Then, if these alternatives are compared in a pair-wise fashion, depending on which pair is compared first, any of the three alternatives could be the final choice.

No matter which method is chosen, fixing costs, effectiveness, or pair-wise comparisons, there will be numerous attendant difficulties and complexities. The best procedure seems to be to pick one method, determine beforehand what seem to be legitimate goals and tradeoffs, and stick with that

method as long as it seems reasonable. There will generally be some point at which informed judgement and preferences for certain outcomes will have to supplement the purely mechanical application of analysis procedures.

3. Differentiating strategy comparable costs and program specific costs: There are essentially two kinds of comparisons which a decision-maker may want to investigate. First, decision-makers may want to compare the absolute costs of various strategies or combinations of strategies in order to identify the general array of all the costs involved in every strategy or combination of interest. If this is the case, strategy comparable costs (adapted from Carpenter, 1978) must be gathered. These costs include virtually everything which is necessary to any of the assessment strategies. There is one exception, and that is where all strategies require the same basic resources, for example, standard classroom space. In this situation, classroom space need not be included, since adding this cost merely inflates the cost of each strategy by a similar amount, and gives an unduly high estimate for all strategies. Resources which can generally be eliminated from costing in this manner are basic resources which are fundamental to the enterprise of education, and not unique to any particular program or activity.

The second type of comparison that decision-makers may want to investigate is one which indicates which strategy or group of strategies is most appropriate for a particular

situation or setting. Program specific costs are appropriate for such comparisons. These costs are costs only for resources which are not already available to a program. Resources which a program already has would not be included in costing of a particular strategy or combination where such resources were necessary. For instance, if a simulation strategy required use of an assessment center, and a center already existed and was accessible, then construction, maintenance, and staffing would not need to be included. Only the costs per use of the center would be figured into the strategy. However, if no such center existed, then, all the costs associated with creating one would have to be included in the strategy cost analysis.

These two different comparison contexts, strategy comparable costs and program specific costs, are both important depending on the level of decision-making involved. If a federal or state agency is estimating average costs for a series of assessment systems, the former method may be more revealing in providing some sense of overall costs. If a specific program is attempting to determine its own system, the latter analysis is obviously of more use.

4. Dealing with non-quantifiable costs and spillovers: In any cost analysis, there will be certain costs which cannot be easily estimated or translated into dollar amounts. In addition, despite the most rigorous planning, there will be unexpected outcomes, often referred to as spillovers, or exter-

nalities, which may produce either positive or negative effects. These are two of the most tenuous areas for the application of empirical analysis in any decision situation; areas where common sense, experience, and intuition must supplement if not supplant more rigorous methodological procedures.

There are some guides which are useful, however. First, when dealing with costs which are not easily quantified, for instance, the costs to students of faulty evaluation which mistakenly eliminates some competent individuals from a program versus the costs to society of faulty evaluation which passes non-competent students into public practice, the various costs should at least be ranked according to their perceived importance. Thus, if it is felt that, given limited resources, society's well-being has relatively more value than students', then a system should be designed which requires more expert performance, and is less lenient in borderline cases, even though no dollar cost is placed on such values. In practice, dollar amounts can often be found for such costs through the measurement of expected utilities of decision-makers or of certain segments of society. Where such criteria are the critical ones of a decision, decision-makers should consult sources such as Raiffa (1968), of the Harvard Business School, Howard and Others, (1977), of SRI, International, and other decision analysis literature.

A second technique for dealing with both non-quantifiable costs and spillovers is to get some estimates of how important

these factors are in the overall decisions to be made. In the above example, one approach might be to determine how much a possible student or public lawsuit might cost an institution if it were determined that the assessment strategies for determining competence were either too stringent or too lenient, and what the probabilities of such judgements might be. Given these estimates, an analysis could be made which would reveal whether different strategies or standards should be used, and at what point additional costs for better assessment ceased to balance the costs to students or society. In other words, if a student lawsuit might cost an institution $10,000, and the likelihood of such a judgement was 1 in 500, then the expected cost would be $10,000 X .002 = $20/ 500 students trained. If the cost of assessment to reduce the likelihood of such an occurrence to 1 in 1000, was $20/student, or $10,000 for 500 students, then the investment would not be reasonable.

Many people are extremely uncomfortable putting questions of value into monetary terms like those above. However, it is essential to realize that putting dollar costs on values such as "fairness to students", or "responsibility to society" occurs at every juncture of the academic process where allocations of scarce funds are made to one type of activity and not another. The argument above is not meant to indicate that student concerns are unimportant, or always of lowest priority. Instead, it emphasizes rather dramatically that tradeoffs are fundamental to

most decisions, and it forces decision-makers to understand that the abstract idea of assessing competence has very real and often contradictory consequences for students and the public.

Although non-quantifiable costs and unexpected outcomes should not be and, in some cases, cannot be ignored, it is generally not worth the investment to attempt to be too refined in placing monetary values on all costs, or anticipating fully all outcomes. Analysis for these concerns is expensive, generally done by outside consultants, and should be undertaken only when the entire program is in jeopardy. More often than not, an awareness of these costs and potential effects is the best that can be hoped for. Unless a cost or an outcome seems so momentous as to outweigh all other considerations, it is simply not feasible to deal with such issues in a detailed analytic fashion.

5. Increasing the validity and reliability of time and usage estimates: The Achilles heel of cost analysis might be seen as the lack of validity and reliability of the various estimates of time and usage which underlie the cost in the Function/Resource matrix for each strategy. There is no magical remedy for this weakness. However, there are some ways to increase the validity and reliability of the estimates. First, clear definitions of functions will help any estimator, but particularly faculty, distinguish activities and so estimate more accurately the proportion of time they spend on each one. Second, the redundancy of estimation in the form of multiple

measures repeated over time, will help identify any inconsistencies and provide some sense of the variance in responses. Third, well conceived and validated measuring instruments and techniques will obviously increase confidence in the final estimates.

Rating scales, interviews, self-report plans and institutional requirements are some of the typical procedures for gaining such information. More sophisticated techniques, for example, some of the procedures used in perceptual psychology, decision analysis, values clarification and research, can be adapted when the costs are significant enough to warrant such investment. Finally, a sensitivity analysis again may reveal that over or under estimates of time may vary widely before any substantial effect on the choice of strategy occurs. (It is quite possible that any systematic over- or under-estimation of time or usage may occur across all strategies and thus, effect all estimates equally.) Again, the only advice which seems reasonable regarding the short-comings of the current methods of cost analysis is that where the analysis seems grossly distorted, then experience, common sense, and intuition should be used to help give a more accurate reflection of true costs.

Primary Cost Analysis

Primary Cost Analysis (PCA) is centered around the Cost Data Matrix (CDM). This matrix is critical to Primary Cost Analysis in two ways. First it guides the analysis, and second, it displays the available data. The Cost Data Matrix is displayed in Figure 4.

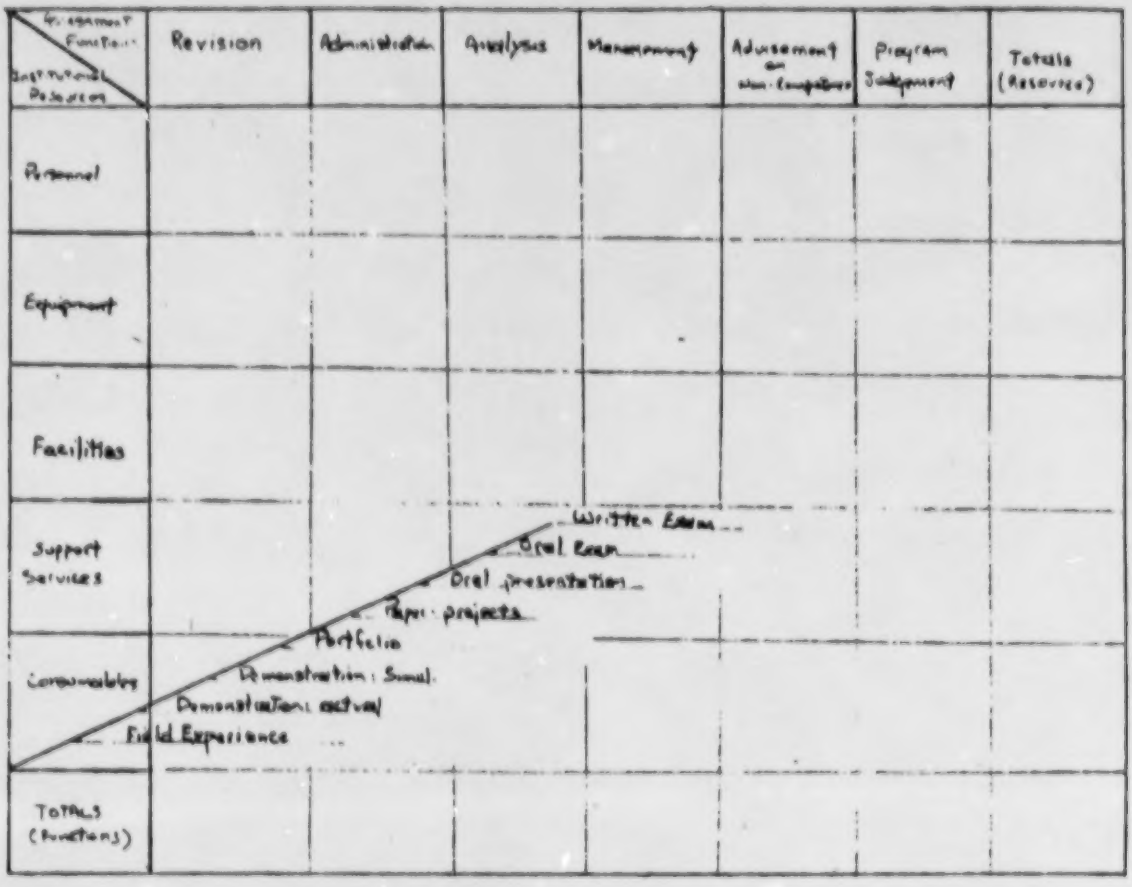


Figure 4 Cost Data Matrix for CBE Assessment Programs

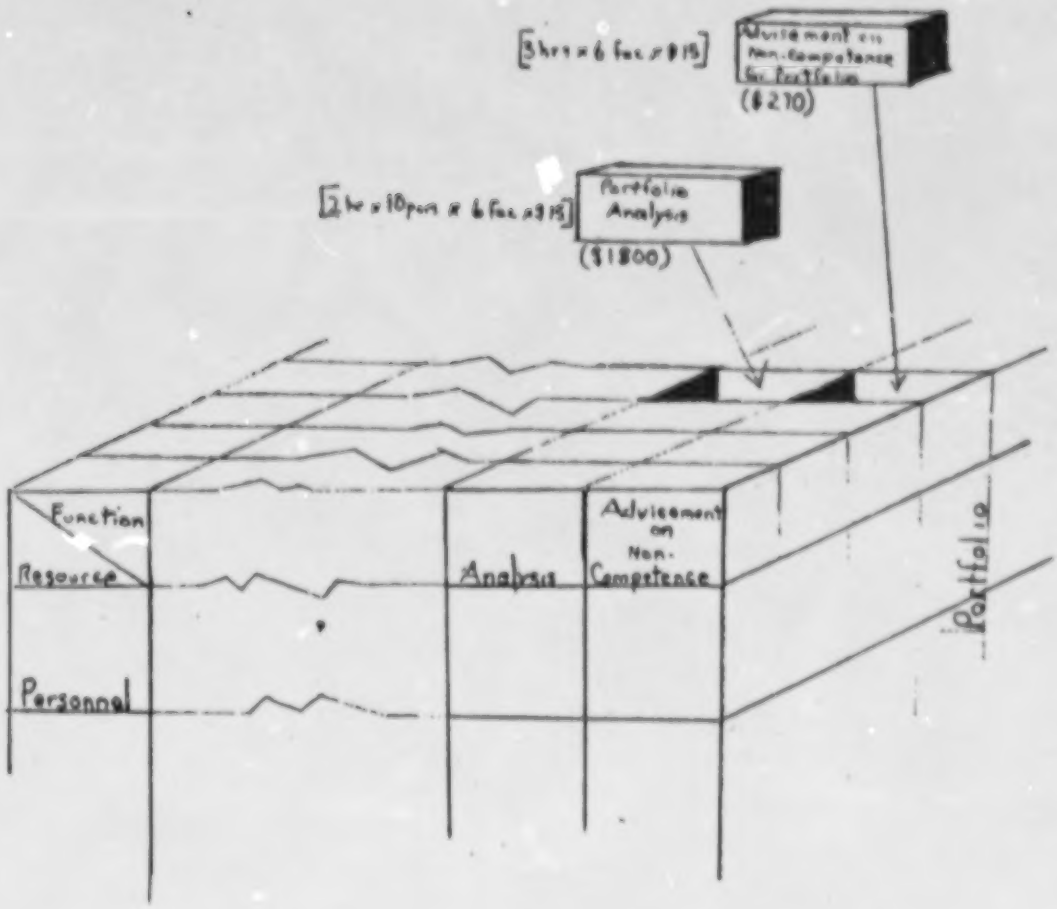
By laying out the units of analysis and their relationship to one another, the matrix serves to inform the evaluator about what to look for in the analysis. The matrix contains three dimensions:

(1) Strategies
(2) Functions
(3) Resources

These are the parameters which determine cost in an assessment system. Each area is sub-divided into several units. There are eight units under Strategies, six units under Functions, and five units under Resources.

The relationship of the dimensions is essentially a function of the interaction of: STRATEGY X FUNCTION X RESOURCE. Each assessment strategy includes certain assessment functions which in turn require certain resources. For example, depending on the procedures, the portfolio (strategy) may require assessment analysis and advisement on non-competence functions by certain faculty members (resources). Once the necessary units of analysis have been identified, it is possible to determine the cost of any given assessment strategy. Figure 3 exemplifies this for portfolio review at Syracuse University.

Figure 3 Hypothetical Cost Analysis for Portfolio Assessment



This illustration shows the total cost of assessing 10 portfolios to be $2,070. That is, each portfolio assessment analysis requires 2 hours from each of six faculty members at a rate of $15/hour, *, or $1,800. Thus analysis costs $180/student. Suppose that for 4 of the 10 portfolios, candidates are being advised on non-competence, and that the format for doing this requires 45 minutes for six faculty members to discuss non-competence with each candidate. At the rate of $15/hour, the cost of the advisement on non-competence is $140, or $67.50/student. These figures represent the hypothetical cost of one assessment strategy which uses two assessment functions and one institutional resource.

---

* The figure of $15 represents the approximate cost to the institution for one hour of instructor time derived by calculating salary plus benefits averaged across the range of ranks and salary levels. The figure is derived from statistical data provided by GRANT and LIND (1978) in which they report the average salary for a 9-10 month full time faculty member in a higher education institution. The figure is also based on data provided by experts at Syracuse University.

Definitions

This section provides definitions for the units in each of the three areas. Understanding these units is critical to the overall analysis.

Assessment Strategies: Assessment strategies are the assessment approaches taken to test student competence. The eight strategies identified by Ely & Daniel (1978) are considered here because they seem to represent the most conventional and feasible methodologies for assessment. In addition, the strategies are further separated into two subdivisions of (a) measures of competence, and (b) indicators of competence, as illustrated below:

(a)	(1)	Field Experience	Measures of Competencies
	(2)	Actual Demonstration	
	(3)	Simulated Demonstration	
(b)	(4)	Portfolio	Indicators of Competencies
	(5)	Paper/Projects	
	(6)	Oral Presentation	
	(7)	Oral Tests	
	(8)	Written Tests	

Assessment strategies as measures are distinguished from strategies as indicators, because, measures assess actual performance in real or simulated job settings, while indicators assess behaviors thought to be more or less related to job performance. The assessment strategies of (1) field experience, (2) actual demonstration, and (3) simulated demonstration are said to measure competence because they are attempting to record the existence of specified, necessary job skills and

attitudes (i.e., the identified competencies). The assessment strategies of (4) portfolio, (5) paper/projects, (6) oral presentations, (7) oral tests, and (8) written tests are said to be indicators of competence because they do not test actual performance either in simulated or actual job situations. Instead, they indicate what kind of performance might be expected. These strategies infer what performance might be like in these situations.

The subdivisions into measures and indicators, then, are based on the following assumptions:

(a) A measure of competence is more valid than an indicator of competence.

(b) Some measures of competence are more valid than others.

(c) Some indicators of competence are more valid than others.

(d) A strong indicator of competence is more useful than a weak measure of competence.

The definitions of the eight assessment strategies are presented below.

(1) Field Experience. Field experience refers to assessment which occurs in the real world context. That is, assessment which occurs in a real job situation or internship. Performance which occurs in this situation is subject to real world effects and tends to demonstrate competence on its most fundamental basis (see Chisholm and Ely, 1976; Hodgkinson and Others, 1975).

(2) Actual Demonstration. Actual demonstration is assessment which occurs in a field setting where many but not all of the real world effects are allowed to influence student performance. Since student access to field settings occur through invitation or through agreement with cooperating institutions, some control exists to limit real world effects.

(3) Simulated Demonstration. A simulated demonstration refers to a microcosm type demonstration. In this case many of the variables affecting performance are controlled or manipulated. Effects due to chance or real world contingencies play only a small part in the assessment situation.

(4) Portfolio. A portfolio is a collection of written materials, audio/visual, artistic, or craft products which are submitted as an indication of field competence. A portfolio is generally composed of materials developed during the course of a program or some professional experience and is used to infer (or demonstrate) job related competence.

(5) Paper/Projects. Normally paper/projects refers to a product emerging from a major assignment. Examples include term papers, paintings, musical compositions, and so forth. Paper/projects may also refer to smaller assignments and projects. The idea here is that the work produced by the student is examined and assessed as an indicator of some knowledge, attitude or skill obtained during instruction.

(6) Oral Presentation. Oral presentations are student-controlled summaries and discussions which reveal (1) the activities involved in, (2) the progress of, and (3) the findings related to student projects or papers. Although oral presentations are more than impromptu conversations, they allow the audience to ask questions or make comments which may require more in depth responses or the extension of ideas. Thus, such testing is more flexible than written presentations, though less detailed.

(7) Oral Test. An oral test is a strategy in which students respond to a set of questions specified by an examiner or group of examiners. This strategy allows flexibility in pursuing answers more fully and clarifying responses, though this is generally at the expense of the organization and detail present in a written response.

(8) Written Test. Written tests are generally of two types: selected response (true-false, multiple choice), and constructed response (short answer, essays). There is virtually no interpersonal interaction in such tests, and competence is assessed on the one-way communication of the student with the written materials.

Assessment Functions. Assessment functions are the discreet activities associated with an assessment strategy. Functions represent what an instructor, supervisor, or credentialing agent actually does in the course of assessing the student.

There are six assessment functions:

(1) Revision (of instruments and procedures)
(2) Administration
(3) Analysis
(4) Management
(5) Advisement on Non-Competence
(6) Aggregate Judgment of Competence

Most people are familiar with the functions of (test) administration and analysis, which as the basis of (test) grading. However, far fewer people, even professional educators, personnel officers, and state accrediting agents are aware of the time necessary to develop, continually revise and improve assessment strategies. Nor are they aware of the time spent in managing the information which comes from the assessment experience. Each of these activities also contribute to the cost of an assessment program and, therefore, must be considered in the analysis. Assessment Functions are defined as follows:

(1) Revision. Revision refers to the adjustments made of an assessment strategy prior to its use. For example, the questions on a test may need revision. The instructor may improve questions, test procedures, or format. Revision also includes the incorporation of new information. For example, at one time in the field of electrical engineering, it became necessary for instructors to include information about transistors. The inclusion of this information required corresponding adjustments in assessment strategies. In fields where high technology or major philosophical shifts require

frequent, drastic revisions, this function becomes a central cost.

(2) Administration. Administration includes the time and procedures necessary to assess students. Traditionally, the standards for this function are the classroom time and activities used to give paper and pencil tests. Administration becomes more complex and costly as higher order assessment strategies are employed. Included in Administration are travel costs.

(3) Analysis. Analysis refers to the activities associated with examining, processing and interpreting performance data from a particular strategy. Judgments of competence on non-competence are based on analysis. The most familiar form of assessment analysis is test correction and scoring. Less familiar forms of analysis are (a) observations related to field assessment, and (b) analysis of paper/projects, portfolios, and oral presentations.

(4) Management. The management function refers to the activities of recording and reporting data and analyses. Normally these activities relate to student grades. However, where instructors approach assessment data with an eye to developing or improving student competence, recording and reporting assessment data become crucial to providing valid diagnosis and prescription. Management also includes planning assessment programs for a semester.

(5) Advisement on Non-Competence. Advisement on non-competence involves all the activities of dealing with students who fail to achieve necessary levels of performance. CBE programs normally recycle these students, but only after advising them of error and misconception. In graduate level courses this function may lead to counselling students out of a program or field and, if appropriate, into another.

(6) Aggregate Judgment of Competence. This function refers to the summative judgment given on a student's cumulative performance at the conclusion of a program. It is the point at which the department, institution, or credentialing agency signs off on overall student competence. The function, as its name suggests, is the aggregation of all the separate assessments into

some single, final decision that a student is or is not ready for professional practice.

Resources. Resources are the entire variety of individuals and materials, (people, supplies, equipment, and facilities) necessary and available for carrying out the assessment functions. There are five resource categories identified for cost analysis of assessment strategies:

(1) Personnel
(2) Equipment
(3) Facilities
(4) Support Services
(5) Consumables

The resources are defined as follows:

(1) Personnel. Personnel are those individuals who function within a discreet instructional unit such as an academic department, or in association with such a unit. Typically they are (1) instructors and field supervisors, (2) managers, and (3) support staff. The critical group consists of the instructors and field supervisors who handle the actual student assessment. Therefore, an analysis of how instructors and field supervisors spend their time on student assessment is critical to a cost analysis of that activity. However, the contributions made to the assessment process by unit managers (e.g. chairpersons) and secretarial and clerical staff must not be overlooked. Thorough analysis will consider the total contribution of all personnel with respect to the assessment process.

(2) Equipment. Equipment includes the durable, reusable hardware and software used in the assessment process. These may be books, computers, games, media equipment, etc. (Equipment is distinguished from consumables because the former is used more than once.)

(3) Facilities. Facilities are classroom, office, and other specialized areas necessary for assessment. When a facility is basic to all assessment strategies, it is not usually included in cost analysis. However, when a specific facility is necessary to only a certain type or class of strategies, it must be costed.

(4) Support Services. Support services refers to those services provided to aid the assessment process. Examples are (1) teaching assistants whose duties consist of administering and analyzing student asessment data, and (2) computer services used to analyze and report assessment data. An important new support service in the assessment program are those services provided by assessment centers.

(5) Consumables. Consumables are those materials which are used during the assessment activity but which are not re-usable at its conclusion. Typically these include answer sheets for paper/pencil tests, record keeping booklets, computer printouts used in assessment, observation forms, etc.

The analysis procedures and definitions presented above constitute what the authors have designated as Primary Cost Analysis. Such analysis involves the direct costing of Strategies X Functions X Resources and the eventual derivation of an actual dollar amount for each assessment strategy or combination of strategies in a given assessment system. It is these dollar figures which are entered in the Cost Data Matrix. The matrix is the final array of the comparable direct dollar costs of the various alternative strategies. This array is the foundation for decision-making based on cost analysis data which can be readily quantified as dollar expenditures.

There exist, however, costs which are less easily translated into dollars, if such a translation is possible at all; for example the value of a human life. In addition, it is often difficult to decide how to distribute costs. For instance, who bears the cost of noise and air pollution from jet aircraft?

Such costs have been previously identified as non-quantifiable (intangible) costs, and unintended outcomes (externalities). These costs are not normally included in the Cost Data Matrix, but nevertheless, must enter into cost analysis. Incorporating non-quantifiable costs and unintended outcomes in the total cost analysis is accomplished through Secondary Cost Analysis, which will be generally defined in the next section. Techniques for secondary analysis have been mentioned earlier, in the section on Non-Quantifiable Costs, and will not be reviewed in the following discussion.

Secondary Cost Analysis

The designation, Secondary Cost Analysis, is a bit misleading. The consideration of non-quantifiable costs and unintended outcomes is secondary only in the sense that the techniques for evaluating such costs and events are much less exact and easy to use. The concerns in secondary analysis are often more significant in and a greater determinant of decisions than the expenditures arrived at in the primary analysis matrix. This is particularly true when secondary analysis involves political and social issues which have ramifications for the public interest beyond the scope of dollars spent. In our previous example, such ramifications might come from CBE when a new program with high standards graduates students with better skills, but produces fewer students. There is a cost to the segment of the

public which might have received some service, though of lesser quality but must now do without any service while waiting for better trained graduates. There are obviously different costs when more graduates are produced, but some are delivering sub-standard service.

Quade (1975) offers a fairly comprehensive list of non-quantifiable costs and externalities. (Externality is the economic term used to designate costs of production which are not borne directly by the producer; thus, it is roughly synonomous with the authors' use of "unexpected outcomes", because most outcomes of consequences affect people who are not directly involved in the production process.) Some of the more fundamental costs for CBE are:

* Social costs
* Opportunity costs
* Accuracy and Information costs
* Inflexibility costs
* Political costs.

Quade does not specifically include the last set of costs, political costs, but these are often the most critical of all in secondary analysis.

Briefly, social costs are those negative external outcomes which affect large segments of the public, in general. Pollution, traffic congestion, nuclear contamination exemplify social costs. Opportunity costs are the costs of the next best alternative for which resources could have been used. This definition is, in fact, the economic definition of "cost" (Fisher, 1967). It

simply means that if the alternatives are to select a performance -based CBE assessment system or a new school gymnasium, the cost of the assessment system is conceived as the cost of the gym - the next best alterntive forgone. Expressing costs as opportunity losses is often much more dramatic than simply presenting a dollar expenditure amount, because it forces people to confront graphically the things they are giving up when they choose one alternative over another.

Accuracy and information costs are central to the discussion of CBE assessment that has so far been presented. Simply, increased accuracy requires increased information which, in turn, entails increased cost. This is the very essence of attempting to cost assessment strategies. Inflexibility costs are also important in CBE assessment, since the commitment to develop a complex system is not easily reversed, nor can much be salvaged if a later decision is made to abandon the project. Inflexibility costs, then, are costs associated with decisions which cannot be easily reversed.

Finally, political costs are those costs which occur generally to decision-makers, and which involve losses in power, status, position, or office. The most obvious case of such costs is one where a politician supports an unpopular law and is not reelected because of it. Some of the political costs of CBE assessment have already been mentioned. In addition, it should be noted that the drive behind CBE is as much political as

educational, and the costs of adopting or not adopting competency programs and assessments should not be ignored in the activity of cost analysis.

Each of these forms of cost resists easy quantitative analysis which would permit an assignment of a dollar value to them. Therefore, they are often overlooked or sidestepped. (One tactic in cost-effectiveness studies is to view such costs as "negative" effectiveness, and let the effectiveness analysts wrestle with the difficulties of measurement). The decision to consider any secondary costs depends on several factors: (1) the knowledge of cost analysis possessed by the decision-maker(s), (2) the availability of time, money and expertise to pursue such information, and (3) the potential impact of such an analysis on the choice among alternatives. As previously stated, for most individual assessment considerations, detailed secondary analysis are generally not feasible. However, a preliminary evaluation of such costs and their potential consequences is minimally necessary.

Combining the Primary and Secondary Analyses

Once both analyses have been completed, assuming a secondary analysis is carried out, some final aggregation of information is necessary. There is no magic formula for such an activity. Instead, decision-makers must consider both sets of data, and weight them according to various program priorities and con-

straints. In some cases, there will be clearly dominant alternatives. That is, one alternative will be better than all others on every criterion. In others, there will be tradeoffs among alternatives and criteria which must be reconciled by the preferences, experience, and intuition of those responsible for decisions.

Summary

The cost analysis model presented above may seem overly complex. This complexity, however, is meant to reflect the scope of potential variables which can affect the cost analysis of a CBE assessment system. The conceptual intricacy which has been outlined above will hopefully not discourage cost analysis efforts, since obviously not all the complications or activities will be present in any particular analysis.

Rather, the numerous issues raised are meant to encourage a thorough cost conceptualization of potential programs, and to discourage the mechanical use of the matrices and procedures developed in this paper. Inadequate conceptualization of the necessary considerations in cost analysis can lead the analyst to commit two extremes of error. Either the analyst may look for data where none exists, or conversely, s/he may overlook data which are essential. Both errors are fatal. There is, then, no substitute for a clear, complete understanding of the alternatives being compared, and of what can reasonably be

expected of the analytic techniques being used.

This last point cannot be over-stated. The authors have attempted to devise a set of procedures which will help people make better decisions about assessment strategies in CBE, by making such decisions more rational. The claim of rationality is based on the assumption that the more information available about the alternatives in a decision, the better the decision will be. Therefore, some cost information, even though incomplete and estimated, should be better than none. Obviously, though, this is not the case when potential or actual cost data seem so incomplete or so erroneous as to be more misleading than helpful. It is the authors' contention that such cases are, in fact, in the minority.

However, a slavish, rote application of the matrices elaborated in this paper would be a disservice to any program. Cost analysis procedures cannot replace human judgment, but rather are one of a number of tools which can aid in the complex educational decisions that people face. If it seems obvious to those most involved in a decision that such procedures are inappropriate, then, it would not be cost-effective to use them!

Given these cautions, the authors maintain that the majority of CBE programs would benefit greatly from both primary and secondary cost analyses of assessment strategies.

Effectiveness Analysis of CBE Assessment Programs

Introduction. In the previous section we discussed the factors critical to determining the costs which operate within a CBE assessment program. These factors are: Strategy X Function X Resource. In this section we will discuss the factors critical to determining the effectiveness of the CBE assessment program.

Effective assessment outcomes are critical to a competency-based program. The believability of the program is based upon effective assessment outcomes. As Harris and Kelly (1976) indicate:

> The credibility of any competency-based curriculum rests directly on the adequacy and appropriateness of its assessment procedures. (p. 247)

According to Harris and Kelly, the effectiveness of an assessment program is illuminated through the determination of "adequacy" and "appropriateness" of assessment outcomes. But, what are adequacy and appropriateness? And, how are they to be determined?

Some attempt to answer these questions will be made in this section. Not specifically however, for the terms adequacy and appropriateness are too vague to assess as such. Rather the discussion will attempt to follow through on the Harris and Kelly suggestion that systematic analysis should occur with respect to the appropriateness and adequacy of an assessment program and

its outcomes.

Determining the effectiveness of CBE assessment outcomes has been a problem in the CBE movement ever since its inception (Reilly and Others, 1977). Kay (1973) asserts that the problems associated with CBE assessment are due to the limitations of the measurement art. McDonald (1974), on the other hand, claims that the problems lie more with a lack of experience with the technology of performance assessment than with any limitation of the measurement art as such. Actually, Kay's argument is part of the broader agrument offered by McDonald. It is likely that much of the current difficulty with developing credible CBE assessment programs is the lack of experience and knowledge with the design and operation of performance assessment. And yet, while the ultimate effectiveness of an assessment program will depend upon the designer's knowledge and technique in assessing and measuring human performance, the effectiveness of that program will be limited in part by the state of the measurement art. For the art of measurement is not without problems and limitations of its own. These problems and limitations, however, are no excuse for avoiding more informed and systematic development of CBE assessment programs or their systematic analysis with respect to effectiveness. This section of the paper offers one useful approach to analyzing the effectiveness of competency-based assessment programs.

Assessment Technology and the Field of Tests and Measurements. Assessment technology is intimately tied to the field of tests and measurements. Hence it is useful to consult this field when developing or evaluating assessment programs. While most educators prefer to leave this esoteric field to the experts, they do so at the risk of developing and operating ineffective assessment programs. For behind the processes of tests and measurement there lies a rich history of theoretical and applied research which can be of considerable help in resolving CBE assessment problems. One competency-based education developer strongly states that the future success of CBE programs will depend upon what CBE educators know about assessment (Holloway, 1978).

A Two Part Analysis. As indicated in the discussion of the general model above, effectiveness analysis is undertaken in two parts: primary effectiveness analysis, and secondary effectiveness analysis. Primary effectiveness analysis deals only with determining the validity of any of the eight assessment strategies (see p. 53). Secondary effectiveness analysis, however, deals with other characteristics of assessment strategies such as strategy format, directions and so forth. Unlike cost cost analysis, effectiveness analysis does not use a matrix to guide the analysis. There is no standard set of variables at play in the major effectiveness analysis. By and large validity alone is considered; and, unless evaluators conduct secondary

analysis, the effectiveness analysis is unidimensional, thereby making the use of a matrix inappropriate. To the degree that evaluators do conduct certain secondary analysis, however, an effectiveness data matrix could be used to display the findings.

Preface to a Discussion of Validation. Hodgkinson and Others (1975) state that competence occurs or fails to occur in real world contexts because it can be inferred that assessment strategies which examine performance may provide the best measures of competence. The closer measures are to the real world contexts, the more valid they tend to be. Strategies which are increasingly removed from real world contexts are potentially less valid as measures of competence. By this reasoning a simulated demonstration would be less valid than a field-based assessment (by how much is not known). Similarly, a portfolio assessment would be less valid than a field-based assessment.

Validity in this work is defined as the extent to which an assessment strategy does the job for which it is used (Mehrens and Lehmann, 1978). When a strategy which has been designed to measure a competency actually does so, it is said to be valid. When an assessment strategy which has only the power to infer a competency is used to measure one, there is need to question its validity. The information resulting from this use of the strategy is of questionable validity. Hence, if the data from a given strategy is to be valid, it must be used for the purpose intended.

Primary Effectiveness Analysis. The outcomes of an assessment program are the results which emerge from each assessment situation. Typically these results are test scores. The effectiveness of these results depends directly upon the effectiveness of the assessment strategy to deliver them. If the strategy is effective, then the results will also be effective.

Effectiveness analysis of assessment strategies is based upon one fundamental question. To what degree does a given assessment strategy do what it is intended to do? In competency-based education this asks to what degree does the assessment strategy show that the student has effectively attained a given competency. Does the strategy provide a valid measure of competency attainment? Attempts to answer this question have been established in the field of test and measurement in which Effectiveness testing is really nothing more than establishing the validity of an assessment strategy as a measure (and/or indicator) of performance. It is probably for this reason that McDonald (1974) recommends the technology of performance assessment.

The major developments in the technology of performance assessment have come at the hands of experts in the field of educational and psychological tests and measurement. It is the concern of these experts to measure some quality, characteristic, skill or competency within an individual and to test the effectiveness of their measures.

The term "test" is usually considered to mean the presentation of a standard set of questions to be answered; the result of which is a score or numerical value which is said to represent a characteristic of a person. This score is also said to be a measure of the characteristic. An IQ score, for example, is a score from a written test which identifies a person's intelligence level. The IQ score is a measure of intelligence.

Because test developers are not always satisfied that the tests they construct actually measure the characteristics they are designed to measure, they take great pains to determine the degree to which the tests actually measure what they are intended to measure. This is called test validation. Rorbach (1971) states that test validation requires a number of activities. Two important ones are discussed here. On is an examination of the test's reliability. The second is a close examination of various types of test validity.

Reliability: A Necessary But Insufficient Condition of a Valid test. Test validation involves a sophisticated analysis of certain test characteristics. Test validation normally begins with an examination of reliability. The reliability of a test is that portion of the result (score) which is due to systematic errors and persists therefore, from sample to sample, or from testing to testing (Kendall and Buckland, 1967). That is, the reliability reflects the degree to which the test can deliver the same or similar scores with each testing, or deliver scores

which reflect the true differences among students. Indications of reliability are presented in the form of coefficients of reliability. For several reasons reliability is a necessary but insufficient condition of test validity. Kerlinger (1973) says

> Reliability, while not the most important facet of measurement is still extremely important....High reliability is no guarantee of good scientific results, but there can be no good scientific results without reliability. (p. 455)

For example, although the test may be reliable, it may be measuring the wrong thing. Reliability measures do not inform one of this fact. It becomes necessary to inquire about such questions through a validity analysis.

Validation: Four Types. Kendall and Buckland (1967) refer to validation as a procedure which provides "evidence that an inquiry is free from bias or otherwise conforms to its declared purpose" (p. 309). In the absence of bias, the inquiry (here we mean test scores) is not influenced by other factors and can be believed. Or where the test results say, for example, that an IQ score is an indication of intelligence and it actually is. Test validity is not easily attained and at best is a matter of degree (Mehrens and Lehmann, 1978).

> A test may have many validities, each dependent upon the specific purpose for which one uses the test. Eventually the validity of any test is dependent upon how it is used in the local situation. (p. 128)

Test validation involves an examination of four types of evidence--evidence related to (1) content, (2) mental constructs, (3) relevant criteria, and (4) appearance. Each of these forms of evidence is discussed briefly below in the manner typically employed by test and measurement experts.

1. Content Validity: The test is examined to determine how adequately the content of the test samples the domain about which inferences are made. An individual's test score can validly infer his/her knowledge of a subject only to the degree that the test has adequately sampled (i.e. provided representative examples of) the subject domain.

2. Construct Validity: The test is examined in terms of the degree to which the test scores can be accounted for by certain constructs of a psychological theory. It is interesting to reflect on how this notion might be replaced in competency-based education by competency validity which would examine the test in terms of the degree to which the scores can be accounted for by certain competencies.

3. Criterion-Related Validity: The test is examined in terms of its relationship with external measures (which measure the same or similar things). For example, success in college (Grade Point Average) is indicated by certain other tests or measures obtained in high school (i.e., SAT, High School GPA). Such an examination speaks essentially to predictive validity. That is, the ability of a test score to predict other types of performance.

4. Face Validity: Here the test is examined in terms of its ability to show how "on the face of it" that it measures what it says it measures.

Test validation is just that, namely, TEST validation, usually in the form of the written test. The technology of

test validation can be applied to OTHER forms of assessment as well as portfolio, oral presentations and so forth. Test validation can be helpful in guiding the validation of all assessment strategies. Because the technology of test validation is a sophisticated technology, it may, however, be necessary to engage experts in early efforts to validate assessment strategies. Persons having some competence in this area may wish to consult the following authors in an effort to gain greater clarity on the technology: Nunnally (1978), Cronbach (1971), and Stanley (1971).

Secondary Effectiveness Analysis. Secondary effectiveness analysis departs directly from the technology of test validation by making inquiry with respect to the following characteristics:

1. Adequacy (Harris and Kelly, 1976).
2. Appropriateness (Harris and Kelly, 1976).
3. Communicability (Doughty and Others, 1976).
4. Equitability (Knapp and Sharon, 1975).
5. Ethicality (Kelly, 1977).
6. Feasibility (Doughty and Others, 1976).
7. Objectivity (Meherens and Lehmann, 1978).
8. Practicality (Meherens and Lehmann, 1978)
9. Utility (Doughty and Others, 1976).

ETC.

Cost is often included in this list. However, because of its importance in the overall C-E analysis, it has been separated into an independant categoryand is not discussed here.

Each of these characteristics pertains to some dimension of an assessment strategy which may influence the choice to keep or remove the strategy from the total assessment program. For example, equitability, as Knapp and Sharon (1975) speak of it, addresses the fairness of the assessment strategy. Scriven (1974) advocates an "inspecting process" which searches for "injustices" within the assessment strategy.

> Whatever the outcomes, there are certain types of procedure (sic) that are inappropriate for moral reasons, and inspection of process must be made to see whether excess cruelty, inequitable methods of grading, etc. arye important. (p. 161)

Clearly where there is evidence that the assessment technique is really not equitable to the student, educators are obliged to discard it, or, alter it until it <u>is</u> equitable. This is particularly difficult in cases where a strategy shows high levels of validity but demonstrates a need for some revision on moral grounds.

The Problems of Secondary Effectiveness Analysis. Two problems are associated with secondary effectiveness analysis. One is the problem of <u>what to consider</u> when conducting secondary analysis. Since there are many characteristics appropriate for

consideration under secondary analysis, some of which may even seem to be conflicting, selecting the appropriate ones can be a problem. This remains a local problem, however, and is to be decided according to the circumstances and conditions under which the analysis is taking place. The second problem relates to definition. Immediate survey of the list above causes one to ask what these various terms mean. What precisely is feasibility or objectivity, for example? Experts are not clear on what these characteristics are, although a sense of these terms is provided within the literature. Nevertheless, the ultimate meaning of any of the terms will be determined again by the location and conditions of analysis and must be dealt with in the course of analysis. Decision makers and analysts will need to decide how secondary analysis applies to their situation. It is hoped that out of such decisions, researchers will develop a better sense of how secondary effectiveness analysis is to be conducted. For the present there is too little information available with which to establish secondary effectiveness guidelines.

The Importance of Secondary Effectiveness Analysis. Secondary Effectiveness Analysis is important because it imposes a realistic perspective on effectiveness-based decisions. There are always more things likely to effect the selection or rejection of an assessment strategy than validity alone. Secondary effectiveness analysis is useful in keeping this fact in view.

Secondary effectiveness analysis also assures a comprehensive analysis. That is, it offers a sense of completeness in the analysis that would otherwise be missing if evaluators confined their analysis to the validation of assessment strategies alone.

Effectiveness Data Matrix. The situational nature of effectiveness analysis prevents a standard matrix approach from being offered from the onset of analysis. As argued above, if an analyst stays only with the primary effectiveness analysis, a matrix is unsuitable for the display of data. However, if the analysis extends into secondary analysis questions, a data display matrix is suitable for arraying available data. Figure 5 illustrates the form such a matrix might take. The matrix arranges the effectiveness data in a way that is (1) complete, (2) shows differences, and (3) is consistent across criteria and strategies. The initial display of data here will inform analysts with respect to differences among strategies--initially in terms of validity, but also in terms of the secondary analysis features selected. For example, it might be that an assessment strategy with good validity shows itself as being infeasibly long. Decision-makers may wish to keep the strategy, but adjust its length; or, they may wish to replace the strategy with one that is equally valid but shorter, thus avoiding the expense of further test development.

Finally, data displayed in the effectiveness matrix is easily transferred to the Data Display for Decision Making. Its orderly arrangement in the effectiveness matrix facilitates transfer to the decision making matrix.

Criteria / Strategy	Primary Analysis		Secondary Analysis			
	Validity	Reliability Coefficients	?	?	?	?
Field Experience						
Actual Demonstrat.						
Simulated Demonst.						
Portfolio						
Paper/Projects						
Oral Presentation						
Oral Exam						
Written Exam						

Figure 5 Effectiveness Data Matrix for CBE Assessment Programs

Conclusion & Discussion

Cost-effectiveness analysis is designed to provide information for deciding between competing alternative solutions to some problem or set of related problems. In this paper the alternative solutions are various assessment strategies. The problem addressed is the identification of a set of strategies (either individually or in combination) that provide the best assessment of competent performance in a CBE library media specialist program within a given level of resources. Cost analysis, as outlined here, begins from two-dimensions of analysis, resources and functions, and incorporates a third dimension of interest to decision-makers, in this case CBE assessment strategies. Thus, the cost matrix includes three dimensions which represent the critical facets of cost which must be considered. That is, this cost analysis examines the alternatives in terms of (1) Resource, (2) Function, and (3) Strategy. The effectiveness analysis matrix, as developed here, is confined to two dimensions, (1) Criteria and (2) Strategy. In the cost analysis, the matrix actually guides the analysis. In the effectiveness analysis, however, the matrix serves primarily to display the primary effectiveness data.

The general model indicates stages of decision making related to the use of various assessment strategies within the CBE context. The first phase in the model is the determination

whether any selection of a strategy or a mix of strategies is minimally feasible based on the cost of the strategies. Second, given that there exists some feasible set, alternative strategies or mixes of strategies are analyzed for the most effective use of available resources. In the early stages of the analysis, each of these concerns is addressed separately - once in the cost analysis and once in the effectiveness analysis. In both cases the scope of analysis is partially dependent upon the amount of secondary analysis which occurs in each area (cost and effectiveness). Thus, data initially is collected and recorded by two distinct processes. Optimal decision-making, however, requires a comprehensive display of cost/effectiveness data. This is made available at the conclusion of the analysis when the information from the earlier stages is brought together for complete viewing in the Data Display for Decision-Making. When the data is displayed in this manner, the decision-maker(s) has a total perspective of the available necessary information (and lack thereof). From this overview, an assessment can be made of the relative utility of any procedure or mix of procedures, using the criteria of cost and effectiveness.

The process of decision-making, however, involves more than simply examining the data display and selecting a procedure. The display also provides some sense of overall effectiveness needed to decide <u>other</u> issues related to a particular program. Specific program goals need to be considered at this point. In the case

of the library media specialist, for example, the major competencies assessed are performance-related. Hence, there should be some format for "performing". In addition, a portion of the program focuses on cognitive knowledge as well. The format for testing this knowledge is traditionally based on paper/pencil exams (good ones, we hope). A decision-maker will need to match techniques to competencies in order to judge which are the most cost-effective techniques.

Furthermore, the use of any particular procedure must be based on an examination of the program - specific matrices associated with each procedure. That is, the decision-maker is advised to return to either the cost data matrix or to the effectiveness data matrix for a detailed examination of the variables influencing the total cost-effectiveness statements in the data display. In the case of costs, this scrutiny will reveal what resources are need and in what amounts. For example, if a particular resource is in short supply or very expensive, then that assessment procedure may be less attractive. Or, it may be that the school or district already possesses the resources; in which case the procedure may cost less than the data display indicates.

Finally, it should be noted that any assessment system will and should contain multiple assessment procedures. The appropriate mix of these procedures, both with respect to (1) the

level of effectiveness need to make valid claims of competence, and (2) the resources needed for a particular program can be estimated from the cost and effectiveness matrices.

The value of any particular cost-effectiveness analysis is based on the worth of the information provided to those responsible for the decision-making. The more accurate the cost and effectiveness data within the matrices, the more valuable is the information for decision-making. The matrices, then, are not in themselves a program panacea for decision-makers, but can help to make the decisions clearer, and thereby increase the potential for arriving at a correct choice.

Future Research

The present work has been conducted largely as a point of departure in the process of introducing systematic cost-effectiveness analysis to procedures for identifying and selecting assessment strategies in CBE. The model presented here is based on general cost-effectiveness methodology, but is a conceptual expension which needs to be refined, clarified, and empirically validated.

Several specific areas which are in need of elaboration are listed below.

1. The specific procedures for collecting and calculating cost data, especially estimating

average time and usage estimates, translating traditional budgets to Function/Resource, and deriving program specific costs, need to be specifically delineated and tested.

2. The relationship of the effectiveness model and test validation procedures needs to be expanded, and validation procedures for non-paper/pencil assessments need to be identified and/or developed.

3. The implied correlations in the Appendix between strategies and degree of effectiveness in assessing competence must be experimentally validated.

Hopefully, this paper has provided a start on these very substantial undertakings.

BIBLIOGRAPHY

BURKE, J.B.; HOUSTON, W.R.; HANSEN, J.H.; JOHNSON, C. Criteria for Describing and Assessing Competency Based Programs, Syracuse, NY: Syracuse University Projects, 1974.

CHISHOLM, M.; ELY, D. Media Personnel in Education, a Competency Approach. Englewood Cliffs, NJ. Prentice Hall, Inc. 1976.

CRONBACH, L. Essentials of Psychological Testing. New York: Harper & Row, Publishers, 1971.

CYRS, T.E.; DOBBERT, D.J. A Competency-Based Curriculum-What Is It? Minneapolis, MN: University of Minnesota, College of Pharmacy, 1976.

DANIEL, E.; ELY, P. Developing Alternative Strategies For Evaluating Competencies of Lib-ary/Media Personnel in Competency Based Programs. Research Proposal, 1978.

DOUGHTY, P.; BEILKEY, A. Cost Analysis and Teacher Education: A Comment on Relevant Relationships and a Review of Existing Models. Albany, NY.: The State Education Department, Division of Teacher Education and Certification, 1974.

DOUGHTY, P.; STAKENAS, P. An Analysis of Cost and Effectiveness of an Individualized Subject Offering. In: C. Sabine (ed) Accountability: Systems Planning in Education. Homewood, ILL: ETC Publications, 1973. Pp. 175-194.

DOUGHTY, P., and Others. Guidelines for Cost-Effectiveness Analysis for Navy Training and Education. San Diego, California: Navy Personnel Research and Development Center, June 1976.

FISHER, G.H. Cost Considerations in Systems Analysis. New York: American Elsevier, 1971.

GRANT, W.V.; LIND, C.G. Digest of Education Statistics, 1977-1978. U.S. Government Printing Office, Washington, D.C. 1978.

GROTELUESCHEN, A.D., and Others. Evaluation in Adult Basic Education: How and Why. Danville, IL.: Interstate Press, 1976.

HAGGART, S.A. The Resource Approach to the Analysis of Educational Project Cost (OE Monograph No. 3, U.S. Department of Health, Education, and Welfare Publication No. 017-080-01914-1). Washington, D.C.: U.S. Government Printing Office, 1978.

HARRIS, J.; KELLY, S.P. Assessment Measures Needed for Competency-based Higher Education. Peabody Journal of Education 53(4), pp. 241-47, July, 1977.

HODGKINSON, N.; HURST, J.; LIVINE, H. Improving and Performance: Evaluation in Higher Education. Berkely, Calif: Center for Research and Development in Higher Education.

KAY, P.M. "Measurement Techniques: What We Have and What We Need". Paper delivered as part of Defining Teaching Competency: Emerging Issues in the Developmental Period of an Assessment Approach to Competency-Based Teacher Education. Symposium at 1973 Annual Meeting AERA.

KELLY, E. Evaluation Techniques, An Overview. Guest Lecture in ITE 631: Concepts and Issues in Evaluation, Gooler, D. Instructor. Syracuse, New York: Syracuse University, Fall, 1977.

KENDALL, M.G.; BUCKLAND, W.Z. A Dictionary of Statistical Terms (2ed), New York: Hafner Publishing Company, 1967.

KERLINGER, FRED N. Foundations of Behavioral Research, 2ed. Holt, Rinehart and Winston, New York, NY, 1973.

Bibliography, continued

KNAPP, J.; SHARON, A. A Comprehension of Assessment Techniques. Princeton, J.J.; Educational Testing Service, 1975.

MANDARES, G.F.; AIRASIAN, P.W. Issues in Evaluating Student Outcome in Competency-based Graduation Programs. Journal of Research and Development in Education 10(3), 1977, pp. 82-92.

McDONALD, F.J. The National Commission on Performance-based Education. Fhi Delta Kappa LV(5), January 1974, pp. 296-298.

MEHRENS, W.A.; LEHMANN, I.J. Measurement and Evaluation in Education and Psychology. Holt, Rinehart and Winston, New York, 1973.

NUNNALLY, J. Psychometric Theory (2ed). New York: McGraw-Hill Book Co., 1978.

PDK National Study Committee on Evaluation. Educational Evaluation and Decision Making. F.E. Peacock Publishers, Inc. Itasca, IL., 1971.

POTTINGER, P.S. Comments and Guidelines for Research in Competency Identification, Definition and Measurement. Syracuse, NY: Syracuse University Research Corp., 1975.

QUADE, E.S. Analysis for Public Decisions. New York: American Elsevier, 1975.

RAIFFA, H. Decision Analysis: Introductory Lectures on Choices Under Uncertainty. Menlo Park, CA: Addison-Wesley, 1068.

READINGS IN DECISION ANALYSIS. Menlo Park, CA.: SRI Decision Analysis Group, 1977.

REILLY, D.H. et al. The Current Status of Competency-based Training. Journal of School Psychology 15(1), 1977. pp. 68-74.

RIKER, W.H.; OREDESHOOK, P.C. An Introduction to Positive Political Theory. New Jersey: Frentice-Hall, 1973.

SCHALOCK. D. "Cost and Benefits". In: From Committment to Practice: The Oregon College of Education, Elementary Teacher Education Program. Oregon: AASTE. In Pre.s. pp. 114-166.

SCRIVEN, M. If the Program is Competency-based, how come the Evaluation is Costing So Much? In: Horston, W.R. (ed). Competency Assessment Research and Evaluation. A Report of a National Conference. Published by The National Dissemination Center, Syracuse, N.Y.: Syracuse University, 1974.

STACE, P. Program Cost Evaluation: An Analysis of the Costs Incurred by Sponsoring a Program of Study Abroad. Unpublished doctoral dissertation, Syracuse University, 1979.